

Earnings Determination:

Contributions from the Occupation and the Establishment

DRAFT
April 6, 2001

Work in Progress

Julia I. Lane

American University, Urban Institute, and U.S. Census Bureau
JLane@ui.urban.org

Laurie A. Salmon

Bureau of Labor Statistics
Salmon_L@bls.gov

James R. Spletzer

Bureau of Labor Statistics
Spletzer_J@bls.gov

We are very grateful to John Abowd, Dan Hamermesh, Bob Lerman, and Mark Loewenstein for their very useful comments on previous drafts of this paper. The views in this paper are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau or the Bureau of Labor Statistics.

I. Introduction

Economists have long known that wages depend on both employee and employer characteristics, as well as the interaction of the two. However, the empirical evidence on the relative importance of individual and establishment effects is limited. This is primarily due to the lack of microdata which links individuals to the establishments where they work, but also due to technical difficulties associated with separating out employee and employer effects. In this paper, we decompose wage variation into employee and employer effects using microdata from the Occupational Employment Statistics program at the Bureau of Labor Statistics. This database contains information from more than half a million establishments, with wages reported for over 34 million individuals in more than 800 occupations.

Our decomposition of wages into employee and employer effects is based on similar work by Groshen (1991b) and Bronars and Famulari (1997). Specifically, we use OLS regressions to partition the sum of squares of wages into worker and establishment components. This decomposition provides information on occupational and establishment wage differentials, the degree of occupational sorting across establishments, the importance of internal labor markets, and the importance of residual individual heterogeneity.

One of the main contributions of this paper is the empirical estimates of how wages are influenced by the establishment at which the individual works. Establishment wage differentials are defined here as the wage premium, controlling for occupation, that is common to all individuals in an establishment. We find employer effects contribute substantially to earnings differences -- the results from our basic model show that controlling for detailed occupation, establishment dummies account for 21 percent of individual wage variation. These employer effects can only be partially explained by observable establishment characteristics such as location, size, age, and industry. These estimated establishment wage differentials, which say that an individual's pay is determined in part by the establishment at which they work, are important for labor economics and theories of the firm.

To better understand our empirical estimates of establishment wage differentials, we examine the correlations of occupational wages within establishments. The theoretical motivation for our correlation analysis is based on team production models, such as Kremer (1993), which predict that workers of similar skill will match together in establishments. The goal of our correlation analysis is to examine the breadth of the establishment wage differentials across occupations. Our results are striking -- we find that establishments that pay well for one occupation also pay well for others. Even after controlling for observable establishment characteristics, we find positive wage correlations for occupations that are closely related as well as for occupations that one would not expect to be closely related in the production process.

We conclude with a discussion of how our results fit into and expand the current literature. Building largely on the work of Groshen (1991a) and Brown and Medoff (1989), we review the various explanations for the existence of employer effects on individual wages. We highlight how our results complement this literature, both with regard to what we know and

what we don't know. We also point out how our empirical work presents new stylized facts to guide future theoretical and empirical work regarding establishment wage differentials.

II. The Wage Decomposition Methodology

Our empirical analysis is based on Groshen (1991b). We have a measure of wages W_{iej} for individual "i" in establishment "e" in occupation "j." We want to decompose the variation in wages into components attributable to occupational differentials, establishment differentials, and differences across individuals. Following Groshen, we estimate the following four regressions:

$$(Occ) \quad W_{iej} = \mu + OCC_j\alpha + \varepsilon_{iej},$$

$$(Est) \quad W_{iej} = \mu + EST_e\beta + \varepsilon_{iej},$$

$$(Main) \quad W_{iej} = \mu + OCC_j\alpha + EST_e\beta + \varepsilon_{iej},$$

$$(Cell) \quad W_{iej} = \mu + OCC_j\alpha + EST_e\beta + (OCC_j*EST_e)\gamma + \varepsilon_{iej}.$$

In these regressions, OCC_j is a vector of dummy variables indicating the occupation, EST_e is a vector of dummy variables indicating the establishment, and (OCC_j*EST_e) is a vector of dummy variables indicating an occupational-establishment job cell.

This wage decomposition partitions the sum of squares of wages into its various components. As Groshen (1991b) mentions, this statistical technique avoids imposing structure on unbalanced data. The OES microdata are unbalanced, with a different number of workers across occupations, and a different number of occupations across establishments. The R-squareds from each of the four regressions are the key to the decomposition. We notationally define these R-squareds as R^2_{Occ} , R^2_{Est} , R^2_{Main} , and R^2_{Cell} .

As seen from the first three regressions above, log wages are regressed on vectors of occupation and establishment indicators separately, and then on both sets of indicators together (the main effects model). The marginal contribution of establishment indicators to the main effects model, relative to the regression with just occupation indicators, measures the portion of wage variation associated unambiguously with the establishment indicators. This is calculated as $(R^2_{Main} - R^2_{Occ})$. Similarly, the marginal contribution of occupation indicators is calculated as $(R^2_{Main} - R^2_{Est})$, and measures the portion of wage variation associated unambiguously with the occupation indicators.

Occupational wage differentials are common in the empirical literature. For example, Groshen (1991b) finds that detailed occupation information accounts for a mean of 20 percent of wage variation in her data. In a survey of establishments such as the OES, with no demographic or labor market information on individual workers, the occupational effects measure wage differences associated with average employee characteristics and skills that differ across occupations. The occupational effects also measure wage premiums resulting

from inherent differences across occupations, such as compensating wage differentials paid to particularly dangerous jobs.

Establishment wage differentials are also common in the empirical literature. Groshen (1991b) estimates that establishment indicators account for a mean of 32 percent of wage variation in her data. Using a similar regression technique with different data, Bronars and Famulari (1997) find that establishment indicators account for 18 percent of wage variation. Establishment wage differentials are the portion of the wage, controlling for occupation, that is determined by an individual's employer. The explanations for the existence of establishment wage differentials is a topic of considerable theoretical and empirical research; we will discuss these explanations later in this paper. In our empirical work, we find establishment wage differentials even after controlling for observable differences across employers.

The explanatory power of occupation and establishment together in the main effects model does not necessarily equal the sum of the marginal contributions to the main effects model from the establishment indicators and from the occupation indicators. This difference, which is measured as $(R^2_{\text{Est}} + R^2_{\text{Occ}} - R^2_{\text{Main}})$, is referred to as the "joint" explanatory power of occupation and establishment. This joint contribution is non-zero if there is any sorting of occupations across establishments. Positive sorting occurs if high wage occupations are concentrated in high wage establishments, whereas negative sorting occurs if high wage occupations are concentrated in low wage establishments. The existing literature -- Groshen (1991b) and Groshen and Levine (1998) -- has found positive sorting between occupational wage differentials and establishment wage differentials.

The interaction of employers and employees has been analyzed quite extensively recently by authors with access to longitudinal linked employer-employee microdata. Using large scale micro data on firms and workers, Abowd, Kramarz, and Margolis (1999) found that individual characteristics contributed about 55 percent to earnings variation in France, while firm characteristics contributed the balance. The interaction effects between workers and firms are substantial -- high-wage workers do sort into high-wage firms. In the state of Washington, person effects accounted for some 50 percent of earnings variation (Abowd, Finer, and Kramarz, 1999; Abowd and Kramarz, 2000), and the covariance between workers and firms range from .0126 (Washington) to .004 (France).

In the fourth regression above, the job cell interactions measure the wage premium paid to a particular occupation in a particular establishment above or below the wage premium predicted by the occupational and the establishment differentials. The relative contribution of the job cells in our wage decomposition is measured as $(R^2_{\text{Cell}} - R^2_{\text{Main}})$. The explanatory power of job cells in a wage regression undoubtedly reflects an employer's compensation policy. For example, the initial phases of an establishment's production process may resemble the average in the industry, but the finishing process may require workers of higher than average ability. Another example may be that entry level workers in a particular establishment are given greater than average training, and are thus paid correspondingly lower initial wages. Groshen and Levine (1998) refer to the relative contribution of the job cells as the "internal (wage) structure effect."

The final contribution to wages is the individual contribution. This is measured as $(1 - R^2_{\text{Cell}})$, and is the portion of the total sum of squares of wages that can not be explained by occupation and establishment indicators. This individual contribution is undoubtedly due to unobserved wage effects from gender, education, tenure, or other individual attributes that are not captured by the interactions of the occupation and establishment indicators.

In sum, this simple decomposition provides information on occupational and establishment wage differentials, the degree of occupational sorting across establishments, the importance of internal labor markets, and the importance of unobserved individual heterogeneity (controlling for occupation and establishment).

III. The Data

We use microdata from the Occupational Employment Statistics (OES) program at the Bureau of Labor Statistics (BLS). The OES is an annual mail survey measuring occupational employment and wage rates by geographic area and by industry. Approximately 400,000 establishments are surveyed each year. Data are collected for the payroll period including the 12th day of October, November, or December, depending upon the industry surveyed. The OES survey covers all full-time and part-time wage and salary workers in nonfarm industries. The survey does not cover the self-employed, owners and partners in unincorporated firms, household workers, or unpaid family workers.

The 1996 survey was the first year that the OES program began collecting wage rate data along with occupational employment data in every State. It should be noted that the OES is not a longitudinal survey. The survey is designed as a three-year sample, with one-third of both the certainty and non-certainty strata sampled each year. The OES microdata have been used by Osburn (2000) for research regarding industry wage differentials.

We use the 1996 and 1997 microdata in our analysis. Our sample has 573,586 establishments with no imputations of wage or employment data. We have occupation and wage information for all of the 34,453,430 individuals employed in these establishments. We also have information on the location, industry, size, and age of each establishment.

The OES survey asks establishments to fill out the elements of a matrix, where occupations are listed on the rows and various wage ranges are listed in the columns. For each occupation, respondents are asked to report the number of employees paid within specific wage intervals. An example of the OES survey form, with many of the occupations omitted for presentation purposes, is given in Figure 1. The OES survey form sent to an establishment contains between 50 and 225 OES occupations. The number of occupations listed on a form depends on the industry classification and size class of the sampled establishments. To reduce paperwork and respondent burden, no survey form contains every OES occupation.

The occupational data in the OES survey are based on the Standard Occupational Classification (SOC) System. Occupations are classified based upon work performed, skills,

education, training, and credentials. There are 824 detailed occupations in our OES microdata. In some of our analysis, we aggregate these 824 detailed (five-digit) occupational codes into seven major (one-digit) occupations: Management, Professional, Sales, Clerical, Services, Agricultural, and Production.

As seen in Figure 1, the wage information provided by establishments in the OES survey is recorded in intervals for either hourly or annual rates of pay. The actual values we use for these intervals are the mean wage of all workers within the interval as computed from the Employment Cost Index for that year.¹ In the following sections of this paper, we discuss the econometrics and the empirical consequences of wage data reported as interval means. All of the wages used in our analysis are measured, in real terms, as the natural logarithm of hourly rates of pay.

The obvious strengths of the OES microdata for economic analysis are the sample size and the level of occupational detail. Specifically, there are more than half a million establishments in our sample, with wages reported for over 34 million individuals in more than 800 occupations. The OES microdata can be viewed as a type of matched employer-employee microdata. Abowd and Kramarz (1999) survey the importance of matched employer-employee datasets towards contributing to our understanding of the relationship between worker earnings and firms. This survey of the literature suggests that there are surprisingly complex interrelationships between workers and firms, and the authors note (page 2704) that "data collected in the future should give information on each job in conjunction with each individual job holder in each individual firm."

This latter criteria cited by Abowd and Kramarz (1999) highlights the potential weakness of our OES microdata. The OES has no demographic characteristics (such as age, race, or gender) or labor market information (such as tenure, experience, or training) for the individual workers. We will return to this point in our discussion of the empirical estimates.

IV. Empirical Wage Decompositions

IVa) Basic Results

We present the results of our wage decomposition in Table 1. In the first column, we report estimates using the seven one-digit occupation measures. In the second column, we report estimates using the 824 five-digit occupation measures. The first four rows report the R-squareds from the regressions described in the previous section. These regressions are estimated from our sample of over 34 million individuals. The bottom five rows report the various contributions of occupation and establishment to wage variation.

The R-squareds in Table 1 demonstrate that knowing an individual's occupation and workplace go a very long way to explaining individual wage variation. More than 72 percent of wage variation is explained by knowing the individual's one-digit occupation and

¹ The interval mean for the bottom interval may vary for states with a higher than national minimum wage. The interval mean for the top interval is set in nominal terms at \$60.01.

establishment, and almost 88 percent of wage variation is explained by knowing the individual's five-digit occupation and establishment. This implies that approximately 12 percent of wage variation is left to unobserved individual heterogeneity (although we acknowledge that this is an underestimate because of our use of interval data).

The importance of the information contained in the detailed occupational categories becomes clear from an analysis of the first row in Table 1. In the first column, the seven one-digit occupation indicators explain more than 28 percent of wage variation. In the second column, the 824 five-digit occupation indicators explain more than 54 percent of wage variation.

The R-squareds in the second row illustrate that establishment indicators alone explain about half of individual wage variation. This regression is of interest other than its intermediary role in our wage decomposition. Kremer and Maskin (1996) develop an index which captures the degree to which workers with similar wages are grouped across establishments. The Kremer and Maskin segregation index is nothing more than the R-squared from a regression of individual wages on a vector of establishment dummies. Our estimate of .4955 is roughly comparable to other estimates from the United States.²

In the bottom half of Table 1, we report the decomposition of individual wage variation into its component parts. Looking at the second column, we find that 25.97 percent of wage variation is associated unambiguously with occupation, and 20.86 percent of wage variation is associated unambiguously with information on the individual's establishment. An important part of the story is the sorting between occupations and establishments -- we find that this joint contribution accounts for 28.69 percent of wage variation. And the final portion of the explained wage variation is the job cell contribution, which accounts for 12.46 percent of wage variation. The residual 12.02 percent of wage variation in the OES data is due to unobserved variation across individuals within a job cell.

It is interesting to compare the results of our wage decomposition with the results reported by Groshen (1991b). If we compute the simple average across the six industries reported by Groshen, her results fall in between the results we report in columns 1 and 2 of Table 1. For example, Groshen's estimates imply that occupation indicators account for a mean of 20 percent of wage variation, and establishment indicators account for a mean of 32 percent of wage variation. Our estimates of the occupation effect range from 15 to 26 percent, and our estimates of the establishment effect range from 21 to 36 percent. Our estimates of the joint sorting effect (14 to 29 percent), the job cell effect (8 to 12 percent), and the individual effect (12 to 27 percent) also compare similarly to the mean of the estimates reported by Groshen (17 percent, 10 percent, and 22 percent, respectively).

² Davis and Haltiwanger (1991) report that 51 to 58 percent of the total variance in wages is accounted for by the dispersion in mean wages across plants. One can manipulate Groshen's (1991b) estimates in her Table 2 and conclude that the R-squareds from regressions of log wages on establishment dummies range from .17 to .86, with a simple mean across the six industries of .504. Bronars and Famulari (1997) report an R-squared of .447. The results in Lane, Lerman, and Stevens (1998) suggest that the proportion of wage variance explained by between firm variation is roughly .45. Outside the United States, Kramarz, Lolliver, and Pelé (1996) report a wage-based segregation measure for France of .38 in 1986 and .48 in 1992, and Bronars, Bingley, Famulari, and Westergaard-Nielsen (1999) report an R-squared of .350 for white collar workers and .455 for blue collar workers in Denmark.

The estimates in Table 1 provide interesting insight into the labor market and the wage setting practices of businesses. The occupation and establishment information in the OES data explain most of the wage variation across individuals. As expected, we find that detailed information on the individual's occupation explains a sizable amount of wage variation. And building on a small but growing literature, we find substantial establishment wage differentials. We also find the sorting of high wage occupations into high wage establishments to be quite important. The empirical evidence also points towards a key role played by internal labor markets, as measured by the job cell contribution to wages.

IVb) OES Wages Measured as Intervals

As mentioned earlier, the OES survey collects employee wage data in intervals. As is evident from Figure 1, this eliminates any wage heterogeneity across individuals within an interval. However, because there are multiple wage intervals for a given occupation on the OES survey form, there is still wage heterogeneity across individuals within job cells (where job cells are defined as an occupation within an establishment). In our decomposition, where we partition the total sum of squares into its components, this interval method of collecting individual wage data should reduce the residual variance attributable to individuals and thus increase the explained variance due to establishments and occupations. How severe might this problem be? In this section, we present an econometric framework for simulating how this data collection methodology affects the estimates from our wage decomposition.

Assume that an individual's true $\ln(\text{wage})$ is Y_{iej} , but the data analyst observes W_{iej} -- the natural logarithm of the OES interval mean. The relationship between the observed wage and the true wage is $W_{iej} = Y_{iej} + \omega_{iej}$, where ω_{iej} measures how the individual's wage differs from the interval mean. For example, in Figure 1, wage interval "H" includes all employees earning between \$19.25 and \$24.24 per hour, and the OES interval mean for survey year 1997 is $W_{iej} = 21.43$. With appropriate transformations to logarithms, ω_{iej} in this example is bounded between $-.1073$ and $.1232$. We shall refer to ω_{iej} as the "interval error."

For any vector of explanatory variables X , the R-squared that we estimate from the regression $W_{iej} = X_{iej}\beta^W + \epsilon_{iej}$ is:

$$R_W^2 = 1 - \left[\frac{(W - X\hat{\beta}^W)'(W - X\hat{\beta}^W)}{(W - \bar{W})'(W - \bar{W})} \right].$$

But the "true" unobserved individual wage (Y_{iej}) should have been used as the dependent variable in the regression, rather than the observed interval mean (W_{iej}). The R-squared that would have been estimated from the regression $Y_{iej} = X_{iej}\beta^Y + \epsilon_{iej}$ is:

$$R_Y^2 = 1 - \left[\frac{(Y - X\hat{\beta}^Y)'(Y - X\hat{\beta}^Y)}{(Y - \bar{Y})'(Y - \bar{Y})} \right]$$

$$= 1 - \left[\frac{(W - \omega - X\hat{\beta}^Y)'(W - \omega - X\hat{\beta}^Y)}{(\{W - \omega\} - \{\bar{W} - \bar{\omega}\})'(\{W - \omega\} - \{\bar{W} - \bar{\omega}\})} \right].$$

If we assume that $X'\omega=0$, so that $\hat{\beta}^Y = \hat{\beta}^W$,³ then

$$R_Y^2 = 1 - \left[\frac{(W - X\hat{\beta}^W)'(W - X\hat{\beta}^W) - 2\omega'W + \omega'\omega}{(W - \bar{W})'(W - \bar{W}) - 2(\omega - \bar{\omega})'(W - \bar{W}) + (\omega - \bar{\omega})'(\omega - \bar{\omega})} \right].$$

If we assume that $W'\omega=0$, and $\bar{\omega}=0$,⁴ then

$$R_Y^2 = 1 - \left[\frac{(W - X\hat{\beta}^W)'(W - X\hat{\beta}^W) + \omega'\omega}{(W - \bar{W})'(W - \bar{W}) + \omega'\omega} \right].$$

Given our assumptions, one can show that $R_Y^2 < R_W^2$. Therefore, when using interval means rather than the true unobserved wages, the R-squareds that we obtain from our regressions overstate the contribution of occupation and establishment indicators to wage variation, and thus understate the residual contribution of unobserved individual heterogeneity.

We should describe our simulation exercise that is based on this econometric framework. We have simulated a $\ln(\text{wage})$ for 34,453,430 individuals from a normal distribution with mean 2.5133 and standard deviation 0.5446 (this mean and standard deviation are reported in the footnotes to Table 1). We then compute the corresponding wage level, and define the interval wage corresponding to the OES intervals reported in Figure 1.⁵ We then define the interval error ω_{iej} as the difference between the individual's true wage and the natural logarithm of the interval wage. And finally, we calculate all the terms necessary to compare R_Y^2 and R_W^2 as defined in the previous section.

We report in Table 1 that the R-squared for the regression of wages on occupation dummies is .5466. If the individual wages instead of interval means were used as the dependent variable

³ This assumption requires some discussion. If X is a matrix of establishment indicators, $X'\omega=0$ implies that the mean of the interval error is zero for every establishment. If X is a matrix of occupation indicators, $X'\omega=0$ implies that the mean of the interval error is zero for every occupation.

⁴ These assumptions are based on the interval wages being the mean of the underlying wage distribution for each interval. Based upon various simulations that we have conducted, we are worried about the validity of these assumptions. One problem may originate if our simulated distribution of wages is not the distribution used to calculate the interval means. A second problem, as noted in footnote 1, is that the interval wage is not the mean for the uppermost interval. We are continuing analysis of these issues.

⁵ For the two open-ended intervals, we follow the OES program and assume that wages less than \$6.00 have a nominal interval estimate of \$4.99 in 1996 and \$5.28 in 1997, and we assume that wages greater than \$60.00 have a nominal interval estimate of \$60.01 in both years. In our simulation, we use the year distribution from the OES microdata and we transform the interval wages from nominal to real.

in the regression, we calculate that this R-squared should be .5234.⁶ Similarly, the R-squared for the regression on establishment dummies should be .4746 instead of the .4955 reported in Table 1. And our simulation suggests that after accounting for the effect of interval means, the R-squared for the main effect regression would fall from .7552 to .7233, and the R-squared for the job cell regression would fall from .8798 to .8426.

Transforming these simulated R-squareds into the occupational and establishment contributions to wage variation, our estimates of {.2597, .2869, .2086, .1246, .1202} reported in Table 1 would change to {.2487, .2747, .1999, .1193, .1574}. Each of the first four terms (the occupational effect, the joint effect, the establishment effect, and the job cell effect) falls slightly, and the residual individual effect rises from .1202 to .1574. Conditional on our assumptions outlined in the previous section (and the concerns expressed in footnote 4 that we are still investigating), we conclude that having individual wage data calculated as interval means does not distort the conclusions we draw from our wage decomposition.

IVc) Establishment Wage Differentials

In column 2 of Table 1, we found that 20.9 percent of wage variation is attributable to differences across establishments. This is strong evidence for establishment wage differentials (EWDs). But before we conclude that EWDs exist, we need to ask ourselves: are these estimated EWDs merely cost of living differences across establishments in different geographical areas? Or even more basic, are these estimated EWDs merely proxying for other characteristics such as size that vary across establishments and are related to wages? Although we may not know precisely why they exist, we do know that employer-size wage differentials exist -- see Brown and Medoff (1989).

Our methodology for controlling for observable establishment characteristics is as follows. Notationally define wages as W_{iejx} , where "x" represents some observable characteristic of the establishment such as geographical area or size. To purge the wage measure of any influence of "x", the wage used in the variance decomposition is $(W_{iejx} - W_x)$, where W_x is the mean of wages calculated over the characteristic x. This "within estimator" is similar to a two-step procedure used by Dickens and Katz (1987) in their study of industry wage differentials.

The wage decompositions after first removing the effects of observable establishment characteristics are presented in Table 2.⁷ The estimates from column 2 of Table 1 are presented in column 1 of Table 2 for ease of comparison. In column 2 of Table 2, we present the wage decomposition after first removing any effects of cost of living differences that are common within counties. The portion of wage variation explained by the establishment indicators falls slightly, from .2086 to .1971. Obviously, local area differences are not explaining why wages vary across establishments. Similarly, in columns 3 through 7 of Table

⁶ The original R-squared of .5466 is calculated as $[1 - (4,633,689/10,219,022)]$. The simulated R-squared of .5234 is calculated as $[1 - (4,633,689 + 451,460)/(10,219,021 + 451,460)]$, where $\omega'\omega = 451,460$.

⁷ We would like to restate that the results presented in Table 2 are from a two-step procedure, rather than the more common method of directly adding control variables to the right hand side of each regression. Since the dependent variable (wages) differs across columns because the first stage is different for each column, we present in the bottom row of Table 2 the standard deviation of wages after removing the effects of observed explanatory variables.

2, we conclude that neither age, size, nor industry controls are singularly explaining why wages vary across establishments. And when we control for all observable effects together in columns 8 and 9 of Table 2, the estimated establishment effect declines but still accounts for between 14 and 18 percent of individual wage variation.⁸ We conclude that establishment wage differentials can only be partially explained by observable establishment characteristics, and thus EWDs are an important explanation for why wages vary across individuals.

An interesting empirical conclusion evident in Table 2 is that the joint effect of occupations and establishments declines dramatically when we remove the joint effects of geography, age, size, and industry. Specifically, the joint effects falls from .2869 in column 1 to .0407 in column 8, and falls further to -.0446 in column 9. We interpret this empirical result as saying that the sorting of high wage occupations into high wage establishments can be largely explained by the observable characteristics of the establishment. Looking at the effects of the observable characteristics one-by-one, industry appears to be the single characteristic that explains the decline of this sorting effect.

We do not find it surprising that the sorting effect declines when we progressively control for more detailed industry. In one sense, detailed industry controls reduce the heterogeneity of occupational mix across establishments. For example, the mix of occupations in a construction firm is much different than the mix of occupations in a law firm, and without controlling for industry we would expect to see high wage occupations sorting together in high wage establishments. But most establishments within an industry use a similar mix of occupations in their production process (albeit to different degrees), and thus we would expect to see much less sorting of occupations and establishments within industry. It strikes us as immediately obvious that geography, size, or age do not reduce the heterogeneity of occupational mix across establishments to the extent that industry does.

V. Occupational Wages Within Establishments

The empirical evidence from our wage decompositions highlights the importance of the establishment for understanding the variation of individual wages. Even after controlling for observable characteristics that vary across establishments, we find substantial evidence of establishment wage differentials. By definition, these establishment wage differentials measure the wage premium paid to all workers in the establishment, regardless of occupation. Although our wage decomposition is flexible enough to allow average wages in particular occupations within the establishment to vary from the establishment average, many of our specifications report estimates of the establishment effect that exceed estimates of the job cell effect. In this section, we further examine these establishment wage differentials by examining the correlations of occupational wages within establishments.

⁸ Note that there are 373,518 unique combinations of county, age, size, and 2-digit industry in the OES microdata, and recall there are 573,586 establishments in the OES data. As is evident from the R-squared from the regression on establishment dummies in column 9 of Table 2, we are getting close to the point of removing all variation across establishments when controlling for observed establishment characteristics. It is for this reason that we do not control for 4-digit industry in the final column.

Our thoughts on this question are guided by the team production model. Kremer (1993) provides a wonderful exposition of this model. The production of goods and services is a multi-stage process, requiring the coordinated and successful completion of distinct tasks. In many production processes, it is not possible for several low skilled workers to substitute for one high skilled worker. As a result, workers of similar skill will match together in firms -- high skilled supervisors will work with high skilled production workers. The empirical implication of this matching process is that we would expect to see a positive correlation of occupational wages within establishments.

Our analysis in this section is similar to previous work by Dickens and Katz (1987), Bronars and Famulari (1997), and Bronars, Bingley, Famulari, and Westergaard-Nielsen (1999). The goal of our correlation analysis is to examine the breadth of the establishment wage differentials across occupations. For example, in a manufacturing plant, we would expect the wages of machinists and production supervisors to be positively correlated since they work side by side on the assembly line. But would we expect the wages of the accountants or the janitors in this manufacturing plant to be positively correlated with the wages of the machinists and the production supervisors?

To better understand our thoughts on this issue, we pursue this specific example. For each establishment in the manufacturing industry, we have computed the mean wage of these four occupations: machinists, production supervisors, accountants, and janitors. In Figure 2, we graph the average wages of one occupation against the average wages of another occupation in the same establishment. There are 338 "data points" in the figure, where each data point represents an establishment.⁹ We find, not surprisingly, that the wages of machinists and the wages of production supervisors are closely correlated (the correlation is .61).¹⁰ We also find that the wages of accountants are positively correlated with the wages of machinists and production supervisors (the correlations are .43 and .41), and the wages of janitors are positively correlated with the wages of machinists and production supervisors (the correlations are .61 and .55). Perhaps most surprisingly, the wages of accountants are highly correlated with the wages of janitors in the same establishment (the correlation is .41).

Although it is outside the scope of our analysis, we would like to mention the enormous wage heterogeneity across the manufacturing establishments that is evident in Figure 2. For example, the establishment mean $\ln(\text{wage})$ of accountants in this sample ranges from 2.1 to 3.9 (with a mean of 2.94 and a standard deviation of 0.26). This heterogeneity is consistent with the findings of Haltiwanger, Lane, and Spletzer (2000), who outline a model where some unobserved business "type" generates heterogeneity in establishment productivity and wages. Furthermore, our findings in Figure 2 of skill complementarity across occupations within the establishment fits quite nicely with Haltiwanger, Lane, and Spletzer's model of complementarity between the "type" of business and the skill composition of its workforce.

⁹ There are 47,633 manufacturing establishments with at least one worker in any of the four occupations. We have selected the 338 manufacturing establishments with at least 2 workers in each of the four occupations.

¹⁰ The correlation coefficient is the square root of the R-squared from an OLS regression of one occupational mean wage against another occupational mean wage. For example, the R-squared from a regression of the mean wages of machinists against the mean wages of production supervisors is .37 ($=.61^2$).

We investigate the relationship of occupational mean wages within establishments more formally in Table 3. For the seven major occupations, we present the correlation matrix of occupational mean wages within establishments. We present two correlations for each occupational pair. The top correlation is unadjusted for observable establishment characteristics, whereas the bottom correlation is based on individual wage data with county, age, size, and 2-digit industry means removed.

Looking at the data unadjusted for establishment characteristics, the average of the 21 off-diagonal correlations is .4614. This is very similar to the estimate of Bronars and Famulari (1997), who report a correlation of mean occupational wages between professional and nonprofessionals of .499. All correlations in Table 3 are statistically greater than zero at conventional levels of significance. One particularly interesting pattern is that all correlations below .4 are in the upper right corner of the table -- it would seem that the least skill matching within establishments occurs between traditional white collar occupations (managers, professionals, and sales) and blue collar occupations (services, agricultural, and production).

Similar to our wage decomposition analysis, we wonder if these correlations are biased upward by not controlling for observable characteristics of the establishment. Looking at the data controlling for observable establishment characteristics, the correlations fall, but each correlation remains statistically greater than zero. The average off-diagonal correlation has fallen dramatically from .4614 to .1618. This leads us to conclude that the occupational mean correlations within establishments are not measuring cost of living differences or establishment size effects, but are measuring the sorting of worker skill by establishments. In other words, establishments that pay well for one occupation also pay well for other occupations. This has interesting implications for theories hoping to explain the source of establishment wage differentials.

```
*****
*** To do: correlation matrix of occupational mean wages within establishments ***
***          using five-digit occupational data          ***
*****
```

VI. Discussion

Using the wage decomposition proposed by Groshen (1991b), we have documented the sorting of high wage occupations into high wage establishments, the magnitude of establishment wage differentials, and the magnitude of internal labor markets -- the wage premium paid to particular occupations in particular establishments above or below the wage premium predicted by the occupational and the establishment differentials. Our basic results show evidence that high wage occupations are concentrated in high wage establishments, although this effect disappears when controlling for industry. All of our specifications show that internal labor markets are important.

One of our key findings in this paper is the large effect that the establishment has on the wages of the individuals who work there. We find that controlling for detailed occupation, 21 percent of wage variation can be explained by information on the individual's establishment. Accounting for observable characteristics of the employer reduces these establishment wage

differentials only slightly. Taking our empirical analysis one step further, we showed that the establishment's wage premium is correlated across all major occupation groups. We now turn to a discussion of how these empirical results fit into and expand the current literature.

Vla) Literature Review

Groshen (1991a) is the classic reference regarding theoretical explanations for establishment wage differentials. She proposes and evaluates five explanations for why individual wages vary among employers.¹¹ The first explanation is that of labor quality, where employers systematically sort workers by ability (as predicted by the team production models). Groshen presents two arguments against the sorting model as the source of establishment wage differentials. First, the differentials are estimated conditional on controls for occupation, and Groshen argues that detailed occupational information proxies quite well for standard human capital variables. Furthermore, industry wage differentials still exist after directly controlling for human capital variables, and still exist even after controlling for unobserved individual ability in a longitudinal analysis. This result is applicable if the process generating industry wage differentials is similar to the process generating establishment wage differentials. Second, it is difficult to reconcile the sorting explanation with findings of establishment wage differentials for all occupations. If sorting occurs because of capital-labor complementarities, or because the production line runs at the speed of its slowest worker, it is difficult to understand why the productivity of janitors and front-office personnel would be affected by the productivity of those on the production line.

A second explanation for the existence of establishment wage differentials is that of compensating differentials. Similar to the first explanation, this is doubtful because compensating differentials such as risk of injury are occupational specific, rather than applying to all workers in the establishment. Groshen (1991a) argues that there is no empirical evidence in the compensating differentials literature that is consistent with establishment wage differentials. Furthermore, the industry wage differentials literature has empirically examined and rejected the hypothesis of compensating differentials.

A third explanation for the existence of establishment wage differentials is that costly information may generate random variation in wages across employers. For example, employers may profit from individuals who find it costly to search for alternative wage offers, or employers who hire infrequently may not have adjusted their pay structure since their last hiring cycle. Groshen (1991a) rejects this explanation based on evidence that employer wage differentials are persistent.

The fourth explanation proposed by Groshen (1991a) for the existence of establishment wage differentials is efficiency wages. Efficiency wage theories, particularly those that emphasize morale, loyalty, and teamwork, can explain why workers in all occupations receive the establishment wage premium. With efficiency wages, heterogeneity across establishments resulting from a variety of factors such as monitoring costs, turnover costs, or managerial tastes generates the heterogeneity necessary to observe establishment specific pay policies.

¹¹ These explanations for establishment wage differentials can also be found in the industry wage differentials literature: see Dickens and Katz (1987), Krueger and Summers (1988), and Katz and Summers (1989).

Unfortunately, there is little, if any, direct empirical evidence on the relationship between efficiency wages and establishment wage differentials.

The fifth explanation proposed by Groshen (1991a) is a model where wage variation across employers results from workers bargaining over rents, or employers sharing profits with employees for other reasons. These models can generate the result that the establishment wage premium covers all occupations. However, the bargaining models are difficult to evaluate. If employers find it profit maximizing to bargain over rents, then Krueger and Summers (1988) would argue that bargaining theories are a variant of efficiency wage theory rather than an alternative explanation for wage differentials. Another difficulty with the bargaining models is their applicability outside the union sector, although the union threat model may apply to workers in the nonunion sector. Groshen finds some support for rent sharing models, citing the empirical literature which tends to show a positive relationship between an individual's wage and the employer's or the industry's profits.

The literature on employer-size wage differentials also provides and evaluates various explanations regarding why the wages of individuals are associated with the establishment where they work. The paper by Brown and Medoff (1989) is the classic reference in this literature. The explanations for the employer-size wage differential offered by Brown and Medoff are very similar to those offered by Groshen (1991a) for the establishment wage differential.

Brown and Medoff (1989) find that compared to the results from a simple wage regression, adding labor quality variables reduces the size coefficients by roughly one-half. Controlling for unobserved labor quality in a longitudinal fixed effects regression, the size coefficients fall by a further five to forty-five percent depending upon the specification. Even so, there remains a significant size effect after controlling for both observed and unobserved labor quality. Brown and Medoff (1989) conclude that compensating differentials, union avoidance, and rent sharing accruing from product market power explain little of the employer-size wage differential. Brown and Medoff present empirical evidence for piece rate workers that leads them to be skeptical of the explanation that large employers pay higher wages because they have difficulty monitoring workers.

Many papers have followed Brown and Medoff (1989) analyzing the employer-size wage differential.¹² A recent paper by Troske (1999) uses linked employer-employee microdata to evaluate explanations that can not be analyzed using most databases. Compared to the simplest regression with just employer-size as explanatory variables, adding individual worker characteristics reduces the employer-size wage premium by roughly 25 percent. Rent-sharing and monitoring are rejected as explanations for the remaining employer-size wage premium. Taking advantage of the linked employer-employee microdata, Troske finds that more skilled workers tend to work together, as predicted by team production models, and this

¹² See Albæk, Arai, Asplund, Barth, and Madsen (1998) for an analysis and evaluation of various explanations using data from the Nordic countries of Denmark, Finland, Norway, and Sweden. They evaluate and reject working conditions, monitoring, and unions as possible explanations for the estimated size effect conditional on standard human capital variables. They also find that the sorting of workers on unobserved characteristics does not explain the size effect.

matching reduces the employer-size wage premium by approximately 20 percent. Controlling for the capital-labor ratio has no noticeable effect on the establishment-size wage premium, but does reduce the firm-size wage premium by 27 percent. After all the data work, Troske concludes that a large and significant employer-size wage premium still exists and remains unexplained.

VIb) Our Contribution

We believe that our empirical results complement the literature just cited. In our wage decomposition, merely knowing the worker's establishment explains 50 percent of the observed wage variation across individuals. Controlling for the seven one-digit occupation indicators lowers this wage variation explained by establishments to 36 percent, and controlling for five-digit occupation indicators lowers this further to 21 percent. If we assume that a worker's detailed occupation proxies for his skills, education, and training, our conclusion that controlling for the worker's occupation explains much, but not all, of the estimated establishment wage differentials is consistent with the above literature.

We have also found that controlling for the observable characteristics of the establishments explains only a small amount of the estimated establishment wage differentials. Controlling for county, age, size, and major industry together reduces the estimated establishment effects from 21 percent to 18 percent, and adding controls for two-digit industry reduces this somewhat further to 14 percent. To the extent that differences in working conditions, union activity, and capital-labor ratios across establishments can be proxied for by observable establishment characteristics such as county, age, size, and industry, we interpret our findings as consistent with the existing literature.

We are now left with the question of how to explain our estimated establishment wage differentials. Any explanation we propose must simultaneously account for our finding that the establishment wage differentials are common to workers in all occupations in the establishment.

Our immediate thought was to consider whether we adequately controlled for differences in labor quality across establishments. We are relying on occupational information to proxy for measures of ability and human capital. While we would like to investigate this directly (but can not with the OES microdata), the work of Groshen (1991b) and Levine (1992) suggests that occupation adequately controls for standard measures of human capital. Furthermore, recent work by O'Shaughnessy, Levine, and Cappelli (2000) finds that measures of skill and job characteristics do not explain much of the difference in wages across employers (although these measures of skill explain quite a bit of wage variation across individuals). Results from these studies suggest that differences in labor quality across establishments are not the leading explanation for our finding of establishment wage differentials.

Our second and related thought was whether we adequately controlled for differences in technology or capital across establishments. Establishment characteristics such as age, size, and especially industry are reasonable attempts at proxying for such differences, but relatively recent work using establishment microdata has illustrated the striking amount of heterogeneity

across establishments within narrowly defined aggregates. It would be useful to incorporate establishment level information on inputs to (and outputs from) the production process into our analysis. However interesting and worthwhile this line of research would be, we need to remind ourselves that differences in technology or capital, by themselves, can probably not produce establishment wage differentials that are common to all occupations.

We believe that any explanation for the existence of establishment wage differentials will rest on a combination of theories. Empirical work from recent analysis of matched employer-employee data shows that higher skilled workers not only work together in the same establishment, but also tend to work with higher quality capital and technology -- see Doms, Dunne, and Troske (1997) and Haltiwanger, Lane, and Spletzer (2000). Modeling these basic human capital results, augmented with a theory of why human resource pay policies might differ across establishments, should show how the gains from skill sorting and capital-labor complementarities can be spread to workers in all occupations in the establishment. These thoughts are not original to us, but run through the existing literature examining why the wages of individuals are affected by their employer. We mention these thoughts in closing as a call for further theoretical and empirical research.

References

- Abowd, John M. and Francis Kramarz (1999). "The Analysis of Labor Markets Using Matched Employer-Employee Data." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, North-Holland Press, pp. 2629-2710.
- Abowd, John M., Francis Kramarz, and David Margolis (1999). "High Wage Workers and High Wage Firms." *Econometrica*, pp. 251-334.
- Abowd, John M., Hampton Finer, and Francis Kramarz (1999). "Individual and Firm Heterogeneity in Compensation: An Analysis of Matched Longitudinal Employer-Employee Data for the State of Washington." In *The Creation and Analysis of Employer-Employee Matched Data*, edited by John C. Haltiwanger, Julia I. Lane, James R. Spletzer, Jules J.M. Theeuwes, and Kenneth R. Troske, North-Holland Press, pp. 3-24.
- Albæk, Karsten, Mahmood Arai, Rita Asplund, Erling Barth, and Erik Støjer Madsen (1998). "Measuring Wage Effects of Plant Size." *Labour Economics*, pp. 425-448.
- Bronars, Stephen G. and Melissa Famulari (1997). "Wage, Tenure, and Wage Growth Variation Within and Across Establishments." *Journal of Labor Economics*, pp. 285-317.
- Bronars, Stephen G., Paul Bingley, Melissa Famulari, and Niels Westergaard-Nielsen (1999). "Employer Wage Differentials in the United States and Denmark." In *The Creation and Analysis of Employer-Employee Matched Data*, edited by John C. Haltiwanger, Julia I. Lane, James R. Spletzer, Jules J.M. Theeuwes, and Kenneth R. Troske, North-Holland Press, pp. 205-229.
- Brown, Charles and James Medoff (1989). "The Employer Size-Wage Effect." *Journal of Political Economy*, pp. 1027-1059.
- Davis, Steve J. and John Haltiwanger (1991). "Wage Dispersion Between and Within U.S. Manufacturing Plants." *Brookings Papers on Economic Activity*, pp. 115-200.
- Dickens, William T. and Lawrence F. Katz (1987). "Inter-Industry Wage Differences and Theories of Wage Determination." NBER Working Paper #2271.
- Doms, Mark, Timothy Dunne, and Kenneth R. Troske (1997). "Workers, Wages, and Technology." *The Quarterly Journal of Economics*, pp. 253-290.
- Groshen, Erica L. (1991a). "Five Reasons Why Wages Vary Among Employers." *Industrial Relations*, pp. 350-381.
- Groshen, Erica L. (1991b). "Sources of Intra-Industry Wage Dispersion: How Much do Employers Matter?" *The Quarterly Journal of Economics*, pp. 869-884.

- Groshen, Erica L. and David I. Levine (1998). "The Rise and Decline (?) of U.S. Internal Labor Markets." Working Paper #9819, Federal Reserve Bank of New York.
- Haltiwanger, John C., Julia I. Lane, and James R. Spletzer (2000). "Wages, Productivity, and the Dynamic Interaction of Businesses and Workers." NBER Working Paper #7994.
- Katz, Lawrence F. and Lawrence H. Summers (1989). "Industry Rents: Evidence and Implications." *Brookings Papers on Economic Activity*, pp. 209-275.
- Kramarz, Francis, Stéfan Lollivier, and Louis-Paul Pelé (1996). "Wage Inequalities and Firm-Specific Compensation Policies in France." *Annales D'Economie Et De Statistique*, pp. 369-386.
- Kremer, Michael (1993). "The O-Ring Theory of Economic Development." *The Quarterly Journal of Economics*, pp. 551-575.
- Kremer, Michael and Eric Maskin (1996). "Wage Inequality and Segregation by Skill." NBER Working Paper #5718.
- Krueger, Alan B. and Lawrence H. Summers (1988). "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica*, pp. 259-294.
- Lane, Julia, Robert I. Lerman, and David Stevens (1998). "Employers, Jobs, and the Dynamics of Earnings Inequality." Unpublished paper, American University.
- Levine, David I. (1992). "Can Wage Increases Pay for Themselves? Tests with a Production Function." *Economic Journal*, pp. 1102-1115.
- Osburn, Jane (2000). "Interindustry Wage Differentials: Patterns and Possible Sources." *Monthly Labor Review*, pp. 34-46.
- O'Shaughnessy, K.C., David I. Levine, and Peter Cappelli (2000). "Changes in Managerial Pay Structures 1986-1992 and Rising Returns to Skill." NBER Working Paper #7730.
- Troske, Kenneth R. (1999). "Evidence on the Employer Size-Wage Premium from Worker-Establishment Matched Data." *The Review of Economics and Statistics*, pp. 15-26.

Figure 1: Example of OES Survey Form
Nonmetallic Minerals and Metal Mining Industries

[illegible]

Table 1: Variance Decomposition

	(1)	(2)
R^2 : $W_{iej} = \text{Occ Dummies}$.2870	.5466
R^2 : $W_{iej} = \text{Est Dummies}$.4955	.4955
R^2 : $W_{iej} = \text{Occ} + \text{Est}$.6468	.7552
R^2 : $W_{iej} = \text{Occ*Est}$.7252	.8798
Occupation	.1513	.2597
Joint Occup & Estab	.1357	.2869
Establishment	.3598	.2086
Job Cell	.0784	.1246
Individual	.2748	.1202
One-Digit Occupation	Yes	
Five-Digit Occupation		Yes

Source: OES unweighted microdata. 34,453,430 individuals.

Wages are measured in natural logarithms: Mean=2.5133, Std.Dev.=0.5446.

There are 7 One-Digit Occupations, 824 Five-Digit Occupations, and 573,586 establishments.

Table 2: Variance Decomposition, Using Five-Digit Occupations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
R^2 : $W_{iej} = \text{Occ Dummies}$.5466	.5359	.5302	.5231	.4831	.4286	.4008	.4436	.4228
R^2 : $W_{iej} = \text{Est Dummies}$.4955	.4497	.4830	.4560	.4206	.3486	.2839	.2173	.0921
R^2 : $W_{iej} = \text{Occ} + \text{Est}$.7552	.7330	.7491	.7360	.7189	.6839	.6526	.6202	.5595
R^2 : $W_{iej} = \text{Occ*Est}$.8798	.8689	.8768	.8704	.8620	.8448	.8294	.8136	.7837
<u>Percentage Contribution</u>									
Occupation	.2597	.2833	.2661	.2800	.2983	.3353	.3687	.4029	.4674
Joint Occup & Estab	.2869	.2526	.2641	.2431	.1848	.0933	.0321	.0407	-.0446
Establishment	.2086	.1971	.2189	.2129	.2358	.2553	.2518	.1766	.1367
Job Cell	.1246	.1359	.1277	.1344	.1431	.1609	.1768	.1934	.2242
Individual	.1202	.1311	.1232	.1296	.1380	.1552	.1706	.1864	.2163
<u>Absolute Contribution</u>									
Occupation	.0770	.0770	.0770	.0770	.0770	.0770	.0770	.0770	.0770
Joint Occup & Estab	.0851	.0687	.0764	.0669	.0477	.0214	.0067	.0078	-.0074
Establishment	.0619	.0536	.0633	.0585	.0609	.0586	.0526	.0338	.0225
Job Cell	.0370	.0369	.0369	.0370	.0370	.0370	.0369	.0370	.0370
Individual	.0357	.0356	.0356	.0356	.0356	.0357	.0356	.0356	.0357
County Controls		Yes						Yes	Yes
Age Controls			Yes					Yes	Yes
Size Controls				Yes				Yes	Yes
Major Industry Controls					Yes			Yes	
2-Digit Industry Controls						Yes			Yes
4-Digit Industry Controls							Yes		
Standard Deviation Wages	.5446	.5214	.5379	.5244	.5082	.4793	.4571	.4372	.4060

Source: OES unweighted microdata. 34,453,430 individuals. Wages are measured in natural logarithms.

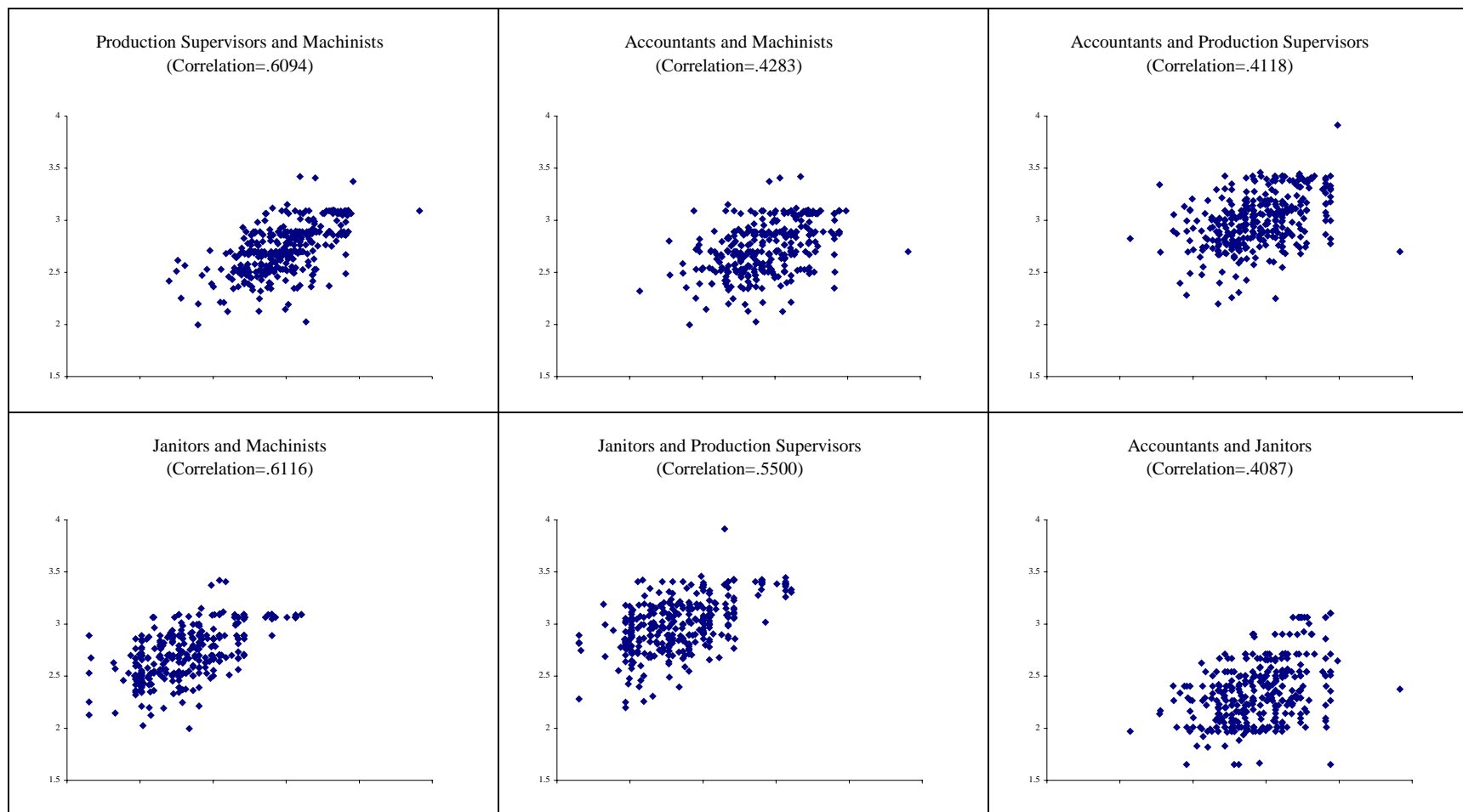
There are 824 Five-Digit Occupations and 573,586 establishments.

There are 3,194 counties, 5 age categories, 9 size categories, 10 major industries, 70 2-digit industries, and 937 4-digit industries.

There are 202,961 unique combinations of county, age, size, and major industry in the microdata.

There are 373,518 unique combinations of county, age, size, and 2-digit industry in the microdata.

Figure 2: Mean Occupational Wages, Manufacturing Industry



Source: OES unweighted microdata. Wages are measured in natural logarithms.
Sample is 338 establishments in the manufacturing industry with at least two employees in each of the following 5-digit occupations: Machinists, Production Supervisors, Accountants, and Janitors.

Table 3: Correlation of Mean One-Digit Occupational Wages Within Establishments

	Management	Professional	Sales	Clerical	Services	Agricultural	Production
Management	1	.5054	.5696	.4503	.3510	.3668	.3790
	1 (N=378,960)	.2978 (N=190,508)	.2508 (N=177,866)	.2710 (N=309,002)	.0971 (N=123,393)	.1212 (N=29,415)	.0354 (N=234,127)
Professional		1	.4515	.4788	.4237	.3625	.4671
		1 (N=242,710)	.1080 (N=95,201)	.2275 (N=212,116)	.0867 (N=91,243)	.0578 (N=20,786)	.0864 (N=126,181)
Sales			1	.5004	.3822	.3869	.5020
			1 (N=263,965)	.0852 (N=179,827)	-.0592 (N=67,313)	.1362 (N=12,940)	.0441 (N=145,992)
Clerical				1	.5138	.4904	.4878
				1 (N=410,387)	.4218 (N=128,401)	.2786 (N=32,757)	.1838 (N=255,165)
Services					1	.5827	.4602
					1 (N=173,193)	.2456 (N=17,470)	.1565 (N=88,471)
Agricultural						1	.5780
						1 (N=41,203)	.2651 (N=25,329)
Production							1
							1 (N=316,958)

Source: OES unweighted microdata. 573,586 establishments. Wages are measured in natural logarithms.

Upper Correlation: No Controls for Establishment Characteristics. Lower Correlation: Controls for County, Age, Size, and 2-Digit Industry.